How to open the door to System 2: Debiasing the Bat-and-Ball problem

Sacha Bourgeois-Gironde, Jean-Baptiste van Der Henst

How to Open the Door to System 2: Debiasing the Bat and Ball Problem

Sacha Bourgeois-Gironde[1] ~~&~~ and Jean-Baptiste Van der Henst[2]

[1] Institut Jean Nicod, Ecole Normale Supérieure, France
[2] Laboratoire Langage Cerveau Cognition, Université de Lyon, France

This chapter presents an empirical study that investigates the conditions under which participants overcome the intuitive and normatively wrong answers they produce in the bat-and-ball problem. In Experiment 1, the problem is phrased in way that highlights its arithmetical constraints. Experiment 2 aims at priming a more analytic procedure by first presenting a similar problem with a less intuitive answer. Experiment 3 introduces a content that invites to evaluate one's own intuitive answer. Experiment 4 manipulates the accessibility of the correct answer in an evaluation task whereas the three previous experiments involve a production task. Only Experiment 4 was successful in eliciting a higher rate of correct answers than in the control condition. Results are discussed with reference to the dual-system approach of reasoning

It is not clear that there is a unified class of cognitive illusions or biases. Wikipedia's entry on cognitive biases proposes a list of about seventy biases affecting judgment, decision-making, and logical or probabilistic reasoning. Error pervades our decisions, reasoning, and judgment processes. The automatic nature of some of our cognitive processes may sometimes lead us astray, making it difficult to recognize and correct subsequent errors. A crucial issue in the assessment of such errors is the clarification of what our systematically biased behavior might count as erroneous. In certain cases sparking heated debates between different research programs on the nature and the extent of cognitive biases, it is unclear that we have a referential norm by which to assess error or, even more problematically, irrationality. One radical premise in the theory of cognitive biases is that there is no genuine error involved. Biases are rather adaptive cognitive strategies given certain circumstances and limited resources of time and cognition. It is inadequate therefore, to label cognitive biases irrational as they could constitute the optimal mental strategies available in conditions of limited cognitive resources.

If this is taken for granted, however, we are left with more clear-cut situations in which error is blatant, regardless of our reluctance to formulate a rational judgment. The controversy remains as to whether a systematically applied procedure leading to a systematic error is a symptom of irrationality. What is not in contention, however, is that in a certain procedure conceived of by the experimenter, the participant may behaviorally deviate from the procedure and an error ensues. This is precisely the case in the arithmetic reasoning task we have chosen to study, which is known as the 'bat-and-ball problem':

> "A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball.
> How much does the ball cost?"

Almost everyone reports an initial tendency to answer '10 cents', as $1.10 is immediately separated between its two parts, integer and decimal—one-dollar-and-ten-cents. This distribution seems intuitive because it is easily processed and because it provides a likely intuitive solution: summing $1 and 10¢ produces the correct amount while an inaccurately processed difference also seems likely, the difference being 90 cents rather than one dollar. and Frederick (2006) report a rate of correct answers, 5 cents, ranging from 17% for participants from University of Toledo to 66% for MIT students.

The bat-and-ball problem is an elementary problem, typical of Luria's stock of arithmetic puzzles, which he studied in order to understand the neural basis of problem solving (Luria, 1966). The problem has recently been used-in an investigation of cognitive impatience or impulsiveness (Frederick, 2005) as well as in mathematical education (Leron, 2006) and legal decision-making (Guthrie, Rachlinski, and Wistrich, 2007). The original empirical data on people's behavior in the bat-and-ball problem were collected by Frederick (2005). This task offers a good opportunity to study the persistence with which an elementary problem typically elicits inadequate procedures and yields an indisputably erroneous answer. It still leaves open the question of whether we should deem irrational the fact that participants employ an inefficient procedure. Typically subjects tend to neglect one of the two operational constraints—sums and subtraction—with which one must comply in order to solve the problem correctly.

However, the diagnosis of irrationality must always be submitted to conditions and is more difficult to pronounce than the assessment of error under a predetermined procedure. It is debatable whether the experimenter is entitled to form expectations about what procedure to follow for a given problem. Once we propose a certain norm and procedure for reasoning, issues of irrationality can be temporarily dismissed and we can proceed without qualms to state whether participants have given an erroneous or correct answer for a certain problem. In the case of problems dealing with simple arithmetic rules, such as that proposed in this study, error and its persistence can be clearly isolated from the more complicated assessment of erroneous participants' irrationality. When dealing with a behavioral issue, one must carefully observe the conditions under which participants persist in their error or correct it, persistence in error being a primary criterion for the presence of a cognitive illusion.

In his Nobel Prize Lecture, Daniel Kahneman used the bat-and-ball problem to present an example of blatant cognitive illusion and illustrate the typical responses from alternative modes of cognitive processing (Kahneman, 2002). Theories of dual reasoning processes

have flourished recently (Chaiken and Trope, 1999; Evans and Over, 1996; Hammond, 1996; Sloman, 1996; 2002; Evans, 2003; Kahneman, 2002; Stanovitch and West, 1998). According to this approach, the typical erroneous answer to which the spontaneous treatment of the bat-and-ball problem gives rise is labeled 'intuitive' and the correct answer is labeled 'normative' or 'analytic.' The main underlying assumption of this approach is that in order to yield a normative answer in a problem of the bat-and-ball type, the mind has to overcome the spontaneous answer. Analytic cognitive processes, aka System 2, have to override intuitive System 1 processes. The bat-and-ball problem is at face value a simple arithmetical problem that any high school graduate could solve. System 2 processes are not in principle challenged by that problem and yet yield widespread error. In that sense, Kahneman is right to emphasize that both System 1 and System 2 are inadequate to deal with the bat-and-ball problem, both System 1 in producing an erroneous answer and System 2 in its failure to correct the error (Kahneman, 2002).

In the current study, our main purpose is to identify the experimental conditions that favor System processes 2 in overriding System 1 processes in solving the bat-and-ball problem. By manipulating the presentation of a task, an experimenter can use hints to facilitate the shift from System 1 to System 2 processing of the task and then observe whether participants change cognitive strategies.

Dual process models come in many guises but all share core assumptions and descriptions such as basic dimensions or characteristic cognitive sub-operations. Kahneman (2002) sums up their typical characteristics as, 1) automatic and intuitive cognitive operations are fast, effortless, and 2) associative, slow learning, generally the product of evolution or of repetitive training, and possibly based on emotions. Controlled reasoning processes are slow, laborious, rule-governed, flexible, and emotionally neutral. The two groups of characteristics apply to sets of cognitive processes distinguished by their speed and controllability, and their operational content. Both types of processes are generally conceived as having different natures. However, one could also imagine that a single cognitive ability (Osman, 2004)—reasoning for instance—can be realized in the mind dualistically, by assuming alternatively different attributes such as, speed, fluency and implicitness on one side, and/or slowness, disfluency and explicitness on the other. This may seem purely verbal but it may have some importance in understanding the question of how System 1 processes are overridden by System 2 processes. Namely, when failing to solve the bat-and-ball problem, should we simply change our reasoning strategies or should we recommence reasoning from the start?

A neutral formulation of the relationship between the two cognitive systems is to assume that System 1 yields intuitive answers to proposed problems and that System 2 monitors the quality of these answers. System 2 monitoring need not be conscious but in case an error is noticed in the way System 1 processes deal with a problem, more explicit representations of the mistaken response and of the correct rules of usage may arise in the participant's mind. In the case of no error being found by System 2, the judgment that System 2 may eventually express is 'intuitive,' in the special sense that it retains the initial System 1 response with no or little modification. System 2 processes may be activated by so-called meta-cognitive experiences such as feelings of difficulty, lack of fluency, and feeling in error. Alter et al.

(2007) have shown that incidental experiences of difficulty—such as making the reading of the task at hand more difficult—tend to reduce the number of intuitive responses proposed for that task.

System 1 and System 2 have been portrayed (Kahneman and Frederick, 2005; Evans, 2003) as competing for control of overt behavior. However, the competition is unequal: System 2 processes are resource demanding in terms of computational ability, memory, and use of operational skills such as calculus and argumentation, and it is normal that System 1 cognitive processes precede and most often preempt System 2 processes. Moreover, intuitive processes, compared to analytic ones, are harder to disrupt (Tversky and Kahneman, 1973; Pashler, 1998). Because they require effort, System 2 processes are easily disrupted by concurrent processes of the same kind. For instance, it is hard to complete two difficult calculations at the same time successfully. One operation hijacks the mental effort and abilities required for the other. By contrast, System 1 processes are not only processed fluidly and easily, but they also form routines and are thus rarely endangered by concurrent processes of the same type. This feature is important when approaching debiasing issues, in that an intuitive but erroneous response to a task is harder to overcome than one resulting from the disruption of System 2 processes. To pursue the issue in the field of arithmetic, it is easier to correct someone who failed to add 189+198 than to correct someone who has falsely learnt and now intuitively believes that 2+2=5.

The greater difficulty in overriding System 1 errors may occur because intuitive responses linger in the mind even after they are replaced by correct responses. Intuitive responsesimpose themselves on the mind in a manner entailing high confidence in their value. It is also possible that even after the participant has realized they are mistaken, they continue to lend credence to the alternate incorrect answer and experience difficulty suppressing it. It is important to emphasize two important criteria of cognitive illusion, high confidence in an erroneous answer, and the resilience of the intuitive erroneous answer. Even though our present goal is not to substantiate directly those proposed criteria of cognitive illusion, we should keep them in mind when trying to understand why it can be so difficult for participants to change their response in the bat-and-ball problem.

As Kahneman and Frederick have clearly stated in the context of a dual-system view, errors of intuitive judgment raise two questions: "What features of system 1 created the error?" and "Why was the error not detected and corrected by system 2?" (see Kahneman and Frederick, 2005, p. 268). In the bat-and-ball problem, it is possible that nobody really bothers to check the validity of one's answer. In other words, System 2's alleged monitoring function is only cursorily, executed, if at all. As asserted by Kahneman and Frederick with respect to the bat-and-ball problem: "People are often content to trust a plausible judgment that quickly comes to mind" (Kahneman and Frederick, 2005, p. 274).

The task is not difficult even though a fringe of the population may be somewhat scared by superficially mathematically formatted problems. However, those individuals should paradoxically perform better on the bat-and-ball problem compared with individuals who take it too easily and promptly generate their erroneous intuitive judgments. In the current study, one focus was to distinguish what it was in the presentation of the task that would divert people from their flawed intuitive strategy. More broadly, how does one secure possible ways out of widespread and robust cognitive illusions that rely on incontrovertible assumptions about which error and correct normative procedures to follow. The choice of a task with inbuilt normativity, due to its arithmetical nature, allows us to feel free to discard some of the aspects of the rationality debate dealing with the cognitive relevance of norms and correct procedures. Instead, focus is directed to a revealing fact with its own lessons as to the nature and extent of human rationality; thus, individuals may be very slow at reaching a correct answer to a task when it contradicts their intuitive inclination.

We added headline numbers. Please check them.

1 Experiment 1: Highlighting the arithmetical constraints

Answering that the ball costs 10¢ intuitively implies that the bat costs $1, because the bat and the ball would cost $1.10 in total as requested by the terms of the problem. Nevertheless, it could also imply that the bat costs $1.10, because the bat would cost $1 more than the ball, as the problem stipulates. The first possibility is wrong because the constraint of $1 difference is not satisfied, and the second possibility is wrong because the sum of the two items amounts to $1.20. It is obviously the first possibility that people are thinking of when they erroneously respond that the ball costs 10¢. Indeed, if participants thought the ball cost $1.10 they would immediately view that this amount already equals the sum of the two items and would obviously understand it is an incorrect possibility. Furthermore, the sum $1.10 is most easily divided into $1 and 10¢. Moreover, any decimal number can easily be seen as a sum of a whole number and a decimal part. Empirical findings confirm that participants think of the '$1 and 10¢' distribution rather than of the '$1.10 and 10¢' distribution. In a pilot study of 38 participants placed in one of Experiment 4's conditions (see below, the Envelope-and-Stamp-production, 50 participants) we asked participants to give the amount of not one but both items of the problem. Only one participant from the pilot study answered that one item cost 1,10 euro and the other 0.10 euro. It is thus clear that when participants answer that the ball costs 10¢ they satisfy the sum constraint but not the difference constraint. One way to prompt people to consider this constraint consists of making it more significant and explaining precisely what it means. If the constraints of the problem become more important, one would expect participants to check more willingly whether their intuitive answers fit. In the experimental condition, both constraints were numbered and presented on separate lines and each constraint was phrased in a more explicit manner than in the control condition. Instead of stating that item one and item two cost 1.10 euro together, Italic it was indicated that the sum of both items' prices–price of item one + price of item two–is equal to 1.10 euro. In the same vein, instead of stating that item one costs 1 euro more than item two does, it was indicated that the difference between the two items—price of item one — price of item two–is equal to 1 euro. In this presentation, we assumed participants would be less likely to overlook one of the two constraints.

1.1 Method

Ninety-eight undergraduate students in social sciences from the University of Lyon partici-

Table 1. Percentage of participants for the three categories of answers in both conditions

	Categories of answers		
	Correct	0.1	Other errors
Control (N=51)	25.5	70.6	3.9
Experimental (N=47)	12.8	70.2	17

pated in this experiment. These 87 female, and 11 male participants were tested in groups of 20-30 individuals. Fifty-one received the problem in the Control Condition and 47 in the experimental condition. All participants received a single problem. The items were not a bat and ball, which are rarely utilized in France, but a stamp and an envelope. We actually referred to a collector's stamp rather than an ordinary stamp, because an ordinary stamp costs 0.55 euro, a price which is much lower than the price expected for the most costly item of the problem (i.e., 1 or 1.05 euro). Hence, in the Control Condition, the problem reads as follows:

> You have bought a collector's stamp and an envelope that cost **1.10 euro** all together.
> The stamp costs **1 euro** more than the envelope.
> How much does the envelope cost?

In the experimental condition, the two constraints were numbered and were more explicit than in the control condition.

You have bought an envelope and a collector's stamp that costs more than the envelope.
12 Given that:

1) The **sum** of the two items, (price of stamp) + (price of the envelope), is equal to **1.10** euro.
2) The **difference** between the two items, (price of stamp) – (price of the envelope), is equal to **1 euro**.

How much does the envelope cost?

1.2 Results and discussion

Answers were classified into three categories: correct answers (the envelope costs 0.05 euro), standard incorrect answers (the envelope costs 0.1 euro), and other incorrect answers (for this category some participants proposed other amounts or wrote that it was not possible to answer). Table 1 shows the percentage of participants in each category. In order to compare the frequency of answers for the two conditions, we used a chi-square test. This was an appropriate test, given that fewer than 20% of the cells had the expected frequency < 5 (see Siegel and Castellan, 1988, p. 123). The chi-square test reveals a significant difference for the conditions (χ^2 (2) = 6.16, $p < .05$). In order to identify where the differences lie, we partitioned the 2 x 3 contingency table into two 2 x 2 subtables and we computed a chi-square

test for each partition. This analysis shows that the difference in correct responses and standard incorrect responses was not significant for the two conditions (χ 2(1) = 1.58, $p < .3$). The analysis also reveals that the Experimental Condition yielded more other incorrect answers than the Control Condition (χ 2(1) = 4.58, $p < .05$).

What conclusions can be drawn from this first experiment? First, the level of performance in the Control Condition was similar to that obtained by Frederick with students attending the University of Toledo (17%; see Kahneman and Frederick, 2006) but was much lower than that obtained from MIT students (66%). Second, the experimental manipulation was not effective. Participants did not produce more correct answers in the Experimental Condition than in the Control Condition; there is even a tendency in the opposite direction. The rates of illusory responses are equally high in both conditions. This means that even if the Experimental Condition prompts participants to pay more attention to both constraints, it does not result in a greater tendency to satisfy the difference constraint. Third, it seems that the formulation adopted for the Experimental Condition tended to make the story even more complicated. While the rate of participants who succumbed to the illusion is identical in both conditions, those that did not succumb to the illusion in the Experimental Condition did not necessarily discover the correct answer and the majority of them actually produced other errors. In the experimental condition, the arithmetic formulation of the constraints might have introduced a source of difficulty and disturbed some social science students who sometimes exhibit poor mathematical skills in France. Thus, in Experiment 2, we abandoned this formulation and explored another debiasing strategy.

2 Experiment 2: priming System 2

As indicated in the introduction, the reason why the bat-and-ball problem shows a low level of performance is not that its solution exceeds the capacities of System 2, but rather that System 1 provides an immediate and effortless answer with a high degree of confidence. Hence, System 2 is not even triggered; if it were, one might anticipate a higher level of performance, but many judgment and reasoning tasks do not always induce an intuitive answer but often require the sole intervention of System 2. Experiment 2 explores the possibility of triggering System 2 in the bat-and-ball problem by first presenting a similar problem that does not prompt an intuitive answer. Since this preliminary problem should involve System 2 and has the same structure as the bat-and-ball problem, one might expect that System 2 will also be triggered for the subsequent bat-and-ball problem. Frederick (2005) reports that the banana-and-bagel problem, which is analogue to the bat-and-ball problem, yielded a higher level of performance: "A banana and a bagel cost 37 cents in total. The banana costs 13 cents more than the bagel. How much does the bagel cost?" The correct answer is 12 cents. There is no obvious solution and one is required to devote more time and effort in order to come up with a solution. In Experiment 2, we used almost the same problem and anticipated that it should give rise to more correct answers than the stamp-and-envelope problem. We then examined whether this improvement in performance affected the subsequent problem.

small point

2.1 Method

One hundred and twenty social sciences students from the University of Lyon took part in this experiment, with 84 female and 36 male participants being individually tested at the Library of Social Sciences. There were 60 participants in the Control Condition and 60 in the Experimental Condition. All participants received the stamp-and-envelope problem; however, in the Experimental Condition, this problem was preceded by the following rubber-and-pencil problem:

> You have bought a rubber and a pencil that together cost **37 cents**. The rubber costs **13 cents** more than the pencil.
> How much does the pencil cost?

In the Experimental Condition, participants were required not to solve the second problem before the first one.

small point

2.2 Results and discussion

We first tested the prediction that the stamp-and-envelope problem is less difficult than the rubber-and-pencil problem. We compared the rate of correct answers for these two problems in the Control and Experimental Conditions and observed that the prediction was confirmed: there were significantly fewer correct answers for the stamp-and-envelope problem (control condition) than for the rubber-and-pencil problem (25% vs. 41.7%, χ 2(1) = 3.75, $p < .06$)[1]. Moreover, in the Experimental Condition there were more participants who only answered the first rubber-and-pencil problem correctly than participants who only answered the second stamp-and-envelope one correctly, (15 participants vs. 5 participants, McNemar χ 2(1) = 4.05, $p < .05$). We then compared the performance for the stamp-and-envelope problem in the two conditions. As shown by Table 2 the distributions of answers are very similar for both conditions (χ 2(2) = 0.9, n.s.).

As with Experiment 1, experimental manipulation in Experiment 2 failed to improve performance. One may argue that the preliminary problem was not entirely appropriate to trigger System 2 because the rate of correct answers was only 16.7% higher than for the stamp-and-envelope problem. However, in a previous pilot study, the same method was followed as in the current experiment but the preliminary problem was much easier . It stipulated that the stamp and the envelope cost 9 euros in total and that the stamp costs 1 euro more than the envelope (hence, the envelope costs 4 euros). Eighty-five percents of the participants correctly solved this problem, whereas only 19% of the participants solved the subsequent problem where the amount of the two items was 1.10 euro (the Control Condition yielded 25% of correct answers). Although participants discovered a successful method that satisfies the sum constraint and the difference constraint when the total amount was 9 euros, there were not able to do so when it was 1.10 euro. Therefore, it seems that even if System 2 is more accessible with a preliminary problem that does not provide a salient intuitive solution, it cannot successfully override System 1 in the stamp-and-envelope problem.

[1] Since there was a straightforward prediction, the level of significance was .1.

Table 2. Percentage of participants for the three categories of answers in both conditions

	Categories of answers		
	Correct	0.1	Other errors
Control (N=60)	25	65	10
Experimental (N=47)	25	63.3	11.7

3 Experiment 3: manipulating the content

The literature on judgment and reasoning has produced a large body of research showing that content and context could drastically modify performance. Two tasks sharing some logical and structural properties but involving different content can result in quite different rates of normative answers. The most eloquent example is certainly the Wason selection task (Wason, 1968) in which the abstract version of the task produces less than 15% of correct answers whereas some thematic versions yield almost 90% of correct answers. The reason advanced for such effects is that content and context activate specific mechanisms that override, interfere with, or facilitate mechanisms elicited by the formal aspects of the problem. For instance, in the case of the selection task some researchers argue that deontic content may activate pragmatic reasoning schemas (Cheng and Holyoak, 1985), specific deontic reasoning skills (Cummins, 1996), or a cheater detection algorithm (Cosmides, 1989) that neutralize formal reasoning. Moreover, in the field of syllogistic reasoning, belief is known to interfere with logical reasoning. The belief bias shows that non-logical conclusions are more widely endorsed when they are believable and that logical conclusions are more widely rejected when they are unbelievable (Evans, Barston and Pollard, 1983). In the field of conditional reasoning, some content activates the retrieval of encyclopedic knowledge, which renders counter-examples to erroneous inferences more accessible (Cummins, Lubart, Alksnis, and Rist, 1991).

In Experiment 3, we wanted to investigate whether System 2 could be stimulated by modifying the content of the bat-and-ball problem. We conjectured that if participants had a greater chance of checking their initial answer they would be more likely to realize that it is incorrect and would try to discover the correct one. We thus constructed a scenario that could encourage participants to verify whether the first response that comes to mind is correct. In this scenario, two founders of a humanitarian organization are fined to 1.1 million euros for embezzling money, but only one of the two founders actually embezzled money.

Consequently, the judge fined the first founder 1 million euros more than the second one. In the control condition, the participant is asked whether he agrees with the fact that the second founder's fine is 100,000 euros. In the three experimental conditions, it is the first founder himself who tells the second one that they have to pay 100,000 euros. Because the first founder is presented as a crook, he may try to cheat the second founder in proposing the '1 million – 100,000 euros' distribution. Thus, the participant may be more cautious about this distribution and may be more likely to check whether it matches the terms of the judgment. In one of the experimental conditions, the participant had to take the perspective of the second founder, the victim of the first founder. We hypothesized that this would enhance the tendency to check the correctness of the distribution since the participant should be more concerned about potential cheating.

3.1 Method

Two hundred and forty-seven psychology students from the University of Lyon participated in this experiment, with 222 female and 25 male participants being tested in two groups of about 120 individuals. Each participant was assigned to one of the five conditions. Because the swindle scenario involves much more information than the standard version of the problem, we wanted to know whether the richness of such a scenario could modify performance. We thus introduced two control conditions: a 'standard control' condition that is similar to the bat-and-ball problem, and a 'swindle control' condition.

3.2 Standard Control Condition

> Someone wants to build a house with a swimming pool. The house and the pool cost 1.1 million euros in total. The house costs one million more than the pool.
> Do you agree with the fact that the pool cost 100,000 euros, and the house 1 million?
> Response: Yes/No (circle your answer)
> If you disagree, indicate how much the pool costs: __________ euros

3.3 Swindle Control Condition

> A few years ago, a humanitarian organization aiming at funding vaccination programs in Africa was created. This organization enabled to collect important amounts of money. However, one founders of the association, Bernard H., embezzled a large part of these funds for his own profit. This swindle was the subject of an investigation and a trial fining the organization has just taken place. The Judge considered that the other founder, as a cofounder of the organization, was also partially responsible for these abuses.
>
> The Judge has fined these two people 1.1 million euros. However, the judge is aware that Bernard H. orchestrated this swindle, and thus fined him 1 million euros more than the cofounder did.

> Do you agree with the fact that the fine of the cofounder amounts to 100,000 euros, and that of Bernard H. to 1 million?
>
> Response: Yes/No (circle your answer)
> If you disagree, indicate how much the cofounder has to pay __________ euros

In the first Experimental Condition, Bernard H. claims that the cofounder has to pay 100,000 euros. The first two paragraphs of the Swindle Control Condition were used and were followed by Bernard H's claim:

> Both people are meeting in order share out the amount of this heavy fine. Bernard H. tells the other person:
> "Your fine amounts to 100,000 euros and mine amounts to 1 million"
>
> Do you agree with Bernard H. about the fact that the fine of the cofounder amounts to 100,000 euros, and his to 1 million?

In the Second Experimental Condition, Bernard H. provides an argument for the '1 million euros – 100,000 euros' distribution. Since Bernard H. is a crook, participants probably expect him to provide an incorrect argument aimed at deceiving the cofounder and would thus be more likely to search for what is wrong with his argument. In this condition, the conclusion of Bernard H's argument omits the difference constraint:

> "Since the judge said that we had pay to 1.1 million euros together and that I had to pay 1 million more than you, you have to pay 100,000 euros and I have to pay 1 million euros, because then the sum is indeed 1.1 million euros."

The Third Experimental Condition is identical to the Second except that the participant has to pretend she/he was the cofounder.

3.4 Results and discussion

All conditions involving the swindle scenario (i.e., the three experimental conditions and the Swindle Control Condition) resulted in unexpected findings. In each condition, about 40–45% of the participants produced amounts that differed from the correct answer and the expected incorrect answer (i.e., 50,000 and 100,000 euros). We think that these participants did not interpret the task in the way intended, that is, as an arithmetical reasoning task, but rather considered that they had to assess whether the sentence was appropriate. For instance, more than half of these participants proposed that the cofounder had to pay 550,000 euros—an amount that never occurs in the standard control condition—probably because they thought that the total fine had to be equally shared between the two people who founded the organization. Others answered that the second founder should pay nothing, probably because that person did not embezzle any money. These unexpected answers might have oc-

Table 3. Percentage of participants for the two categories of answers in the 5 conditions

	Types of answers	
	Correct	0.1
Standard control (N=45)	17.8	82.2
Swindle control (N=25)	20	80
1st Experimental Condition (N=25)	20	80
2nd Experimental Condition (N=30)	26.7	73.3
3rd Experimental Condition (N=33)	18.2	81.2

curred because of the ambiguity in the question's wording. It actually asks the participant whether they agree with the fact that the fine of the cofounder amounts to 100,000 euros, In the swindle scenario, this may be understood as whether this amount is fair with respect to the illegal action raised by the scenario rather than whether it is arithmetically correct. Given that we were not interested in this alternative interpretation, we focused our analysis only on the correct answer and the expected incorrect answer for the five conditions. We thus discarded 10 % of the participants in the standard control condition, 46% in the Swindle Control Condition, and 47%, 41% and 38%, respectively, in the three experimental conditions. Table 3 shows the distribution of answers between the two categories of answers for the five conditions.

Italic

A chi-square test shows that these distributions do not differ significantly between the conditions (χ 2(4) = 1.02, n.s.). Once again, the experimental manipulation failed to boost the rate of correct answers. The content of the swindle scenario did not enable the successful contribution of System 2. Although the person who proposed the '1 million–100,000 euros' distribution is presented as a crook, participants could not find a way to correct this intuitive distribution supporting the correct response.

So far, the debiasing methods applied, were aimed to induce participants to check their mistaken intuition, but all failed. One could of course try to improve such methods by making them more compelling but one could also consider that overall the 'checking' approach is ineffective. It actually rests on the idea that the correct answer is within the scope of System 2 but if this assumption is false, debiasing is most likely impossible. Participants might lose faith in their intuitive answer but given that they cannot not find any satisfactory alternative they are likely to stick with it. In Experiment 4, we tried to make the correct answer more accessible rather than only prompting participants to inspect their intuitive response.

4 Experiment 4: Making the correct response accessible

In Experiment 4, we manipulated two variables, the number of items concerned by the question, and the accessible of the correct answer. First, when the question is about two items instead of just one (the price of the stamp and the price of the envelope vs. the price of the envelope only), it is easier to see whether the sum and the difference constraints are satisfied.

As in the previous experiments, this might help participants to notice that the intuitive answer is incorrect. Second, if participants were not able to consider the alternative correct answer by themselves, making that answer accessible might help them see that it is actually correct. In two conditions, three answers were proposed—the correct answer, the intuitive answer, and another incorrect answer—and participants had to decide which one was correct. The combination of these two hints should lead people to consider a correct alternative and the reason why it is correct.

4.1 Method

EN-Dash

One hundred and eighty-nine students majoring in Psychology and Linguistics at the University of Lyon participated in this experiment. Because of a mistake in the preparation of the material participants' gender was not available, although most of them were female. The experiment employed a 2 × 2 design with two types of questions: 1) a question about the envelope only vs. question about the stamp and the envelope) and 2) two levels of accessebility in the correct response: an evaluation task that proposes the correct answer with two incorrect answers vs. a production task. Each participant was assigned to one-of-the-four conditions. The Envelope-Production Condition involved exactly the same stamp-and-envelope problem as in the Control Conditions of Experiments 1 and 2. In the Envelope-Evaluation condition, the question was followed by three possible responses that could be presented in two different orders: '0.10 euro,' '0.05 euro,' '0.0 euro' vs. '0.05 euro,' '0.10 euro,' '0.0 euro.' In the Envelope-and-Stamp-Production condition, the question was asked, "How much do the stamp and envelope each cost?" In the Envelope-and-Stamp-Evaluation condition, the question concerned both items and was followed by three possible responses that could be presented in two different orders, '1 euro and 0.10 euro,' '1.05 euro and 0.05 euro,' '1.10 euro, and 0.10 euro' vs. '1.05 euro and 0.05 euro,' '1 euro and 0.1 euro,' and 1.10 euro and 0.10 euro.'

4.2 Results and discussion

Figure 1 presents the percentages of correct and incorrect answers in the four conditions. A chi-square test reveals that the rates of correct and incorrect answers differ between the conditions (χ 2(3) = 10.4, $p < .02$). In order to identify where the differences lie, we partitioned the 4 x 2 contingency table into three 2 x 2 subtables and a chi-square test was computed for each partition. This analysis revealed that there were 1) no significant differences between the Envelope-Production and Envelope-Evaluation conditions (χ 2(1) = 0.4, n.s.), 2) no significance differences between these two conditions and the Envelope-and-Stamp-Production condition (χ 2(1) = 0.9, n.s.), and 3) more correct answers in the Envelope-and-Stamp-Evaluation condition than in the other three conditions (χ 2(1) = 9.07, $p < .005$).

In contrast with the previous experiments, the current experiment included a condition that significantly improved performance. This condition combined two hints, the question concerned both items instead of only one and participants had to evaluate, rather than produce, the correct answer. The first hint should have helped participants to check whether

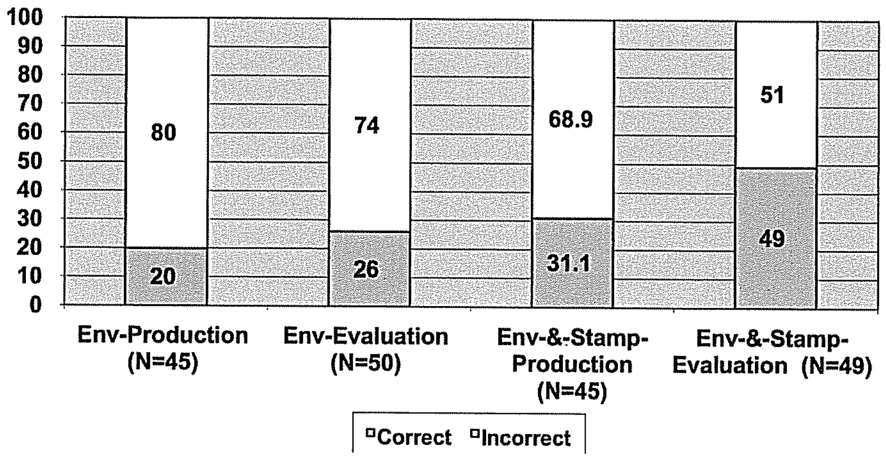

Figure 1
Center

their intuitive answer was correct and the second hint makes the correct response more accessible. Interestingly, it is only when these two hints were presented together that performance increased significantly. We already knew from the previous experiments that prompting participants to check their answer had not really helped them find the correct response. The current experiment also showed that presenting the correct answer was not sufficient to improve performance. This seems to indicate that in the Envelope-Evaluation condition, participants paid attention only to the proposal that matched their intuitive answer and did not really consider other responses. It is thus only when the correctness of the proposed answers can be very easily checked that a significant part of the participants escape the illusion.

General discussion

Kahneman (2002) argued that errors of intuitive judgment, such as the one that occurs in the bat-and-ball problem, have a two-fold origin. First, System 1 yields the error and then System 2 does not detect it or does not correct it. We presumed that the accurate answer to this problem was within the reach of the university students who participated in our experiments. More precisely, this implies that once participants detect a mistake in their intuitive answer, they should be able to correct it. In order to detect an error, one must of course check if the solution that initially comes to mind conforms to the terms of the problem it addresses. For the bat-and-ball problem, the answer has to satisfy a sum constraint (the amount of both items is $1.10) plus a difference constraint (one item costs $1 more than the other does). We developed an experimental procedure that aimed at eliciting System 2 processes by enhancing the tendency to check the conformity of one's answer with the terms of the problems. In Experiment 1, we highlighted the constraints of the problem that had to be respected. In Experiment 2, we presented a problem with no intuitive answer so that participants would

more likely focus on the arithmetical constraints of the problem and better focus on a subsequent problem. In Experiment 3, we manipulated the content of the problem so that participants should become more vigilant about the distribution they had to evaluate. None of these debiasing attempts was successful.

Debiasing strategies aim to make people give a normative answer in a task in which they initially perform poorly. Some of these strategies tend to make people realize the procedures they are using are leading them to an incorrect answer to the problem and others–independently of meta-cognitive awareness attained by participants–simply aim at behavioral correction. Fischoff (1982) lists the extant methods of debiasing that have been used in connection with two well-documented cognitive biases, the hindsight bias, and overconfidence. He describes categories under which it is possible to classify debiasing techniques according to the source to which the bias is attributed: the nature of the task, the misunderstanding of the task, a cognitive default in the participant that may be corrigible or not, and, finally, the incompatibility between participants and the task. In fact, an experimenter must have an idea of where the source of the bias they study resides in order to select a debiasing technique.

In this case, one might hesitate to attribute the source of the bias to the mode under which the task is presented. Nevertheless, if one suspects that the initial version of the task will be poorly treated because of a misleading presentation or because of misunderstanding on the part of the participant, the rephrasing suggested in Experiment 1 should dissipate those risks. The requirements to solve the problem were made explicit from an arithmetical point of view.

Likewise, if the source of the bias is attributed to the lack of ecology in the task–nobody in fact really projects their self into the post office buying a stamp and an envelope for the amounts indicated in our problem. One can argue that our Experiment 3 put participants in a more realistic situation by forcing interest on the part of the subject for the contents of the stories.

If not the task, then participants should be incriminated. However, the situation here was not that clear-cut. The hypotheses have been phrased to guiding the study in terms of a competition between System 1 and System 2. In the first three experiments, it seems that System 2 failed to take over. Were our participants unable to detect their error? For some participants, the answer was clearly yes, as is paradoxically shown by the results of Experiment 4. Although the Envelope-and-Stamp Condition of this experiment enabled an extremely simple detection of the mistake as well as the correct answer, about half of the participants still succumbed to illusion. Certainly, this Evaluation Condition is the one that produces the highest rate of correct answers but the fact that such a high proportion of participants still err in this Condition tells a great deal about the strength of the illusion. Thus in the first three experiments, some participants were certainly insensitive to experimental manipulation and did not notice errors in their answers. However, the absence of performance improvement in these experiments does not necessarily imply that all of the participants failed to consider something was wrong with their intuitive answer. If participants detect the error without being able to discover a way to solve the problem, they may stick with the initial intuitive answer. This contradicts our presupposition that detection should imply correction in the bat-and-ball problem and that the procedure required to solve the problem is within the scope of System 2 processes. Nevertheless, when a structurally identical problem is provided, performance definitely increased but in quite a modest proportion (in Experiment 2 about 60% of the participants failed to solve the rubber-and-pencil problem). Moreover, it might be the case that the presence of the illusion mobilized too much attention and somehow inhibited correction processes even when participants were aware of an error in the problem. In the future, it may be worth investigating System 2 processes as a combination of both error detection processes and correction processes. In order to see whether the procedures introduced in the first three experiments had an effect on error detection, it could be worth adding measures of confidence. If participants report a lesser degree of confidence in the experimental conditions than in the control one, this will be a sign that error is being detected or that at least a feeling of error is being experienced.

How should one interpret the results of Experiment 4? Is the manipulation that consisted in presenting possible amounts for both items a genuine debiasing method? One could argue that debiasing was successful since performance significantly improved in this condition, with more than twice as many participants being right, as opposed to the control condition. Nevertheless, the task was radically different, and the type of System 2 processes required to deal with the evaluation task need not have been as sophisticated as those required to solve the production task. In the production task, participants had to develop a subtle method in order to reach a distribution that conformed to both arithmetical constraints whereas in the evaluation task participants simply had to check whether the proposed answers satisfied those constraints. In other words, detection processes were more easily triggered in this experiment but one cannot say that a successful implementation of corrective procedures occurred.

Finally, the current study sheds some light on the nature and extent of cognitive illusions. Unlike some, or even most cognitive illusions that have been listed in the heuristics-and-biases literature (Tversky and Kahneman, 1973), the bat-and-ball problem is not controversial enough to the presence of a genuine error. Arithmetic norms are rarely put into question even by fervent critics of rationality. Of course, one can dispute the accessibility of the problem, but it is in fact a simple problem and the availability of the correct answer was purposely upgraded across successive manipulations.

This problem lends itself particularly well to an approach in the terms of a theory of dual processes of reasoning and proposed in part as an account of cognitive illusions. We have simply characterized cognitive illusions as robust cognitive biases or errors in the treatment of reasoning tasks. More defining features that we would like to advance include their ubiquity, for most subjects fall into them, robustness and, or resistance, that is difficulty in overcoming, and their possible resilience. Like their cousins perceptual illusions, they tend to recur even after one becomes aware of them. The research program that has most closely approximated the possibility of criteria and classification of cognitive illusions and drawn an explicit parallel between perceptual and cognitive illusions is clearly the heuristics-and-bias program. Sources of cognitive illusions are then found in the use of the basic heuristics of availability, representativeness, and anchoring (Tversky and Kahneman, 1973). Tversky and Kahneman's underlying fundamental hypothesis states that cognitive illusions share features with perceptual illusions in the sense that outputs of judgment and reasoning processes occur as if perceived by the mind and as such are submitted to biases analogous to biases in perception. In the current study, the first two features were established. It will take more study to ascertain the resilience factor and its ability to explain the failure of System 2s monitoring of the task-treatment to yield implementation of corrective procedures in the bat-and-ball problem. However, we can tentatively hypothesize that the kinship of cognitive and perceptual illusions is not fortuitous and is more than an analogy from an evolutionary point of view. In order to implement the 'fast and frugal' heuristics (Gigerenzer, Todd, & the ABC Research Group, 1999) of mind/brain cognitive resources typically dedicated to the automatic processing of information, modules of perception are recycled in the treatment of reasoning tasks that, normally, would require slower and more controlled processing. Most of the time, this rapid treatment is satisfactory, but as it is the case with the bat-and-ball problem, the 'perceptual' treatment of the task yields a deficient solution. The more the treatment of a task is encapsulated in an automatic treatment, the harder it is to overcome by System 2 cognitive processes.

Acknowledgments

We thank Mathilde Bonnefond, Pierre Stanislas Grialou, Hugo Mercier, Ira Noveck, Mariia Kaliuzhna and Nadia Sachs for their help in the preparation of the work presented here. This study was supported by a grant from the Agence Nationale pour la Recherche (ANR BLAN07-1_191653)

References

We arranged the style of References according to the Instructions, which please note.

Alter, A. L., Oppenheimer, D. M., Epley, N., and Eyre, R. N., 2007. Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General* **Vol. 136** (4), 569-576. Roman

Chaiken S. and trope Y. (Eds.), 1999. *Dual-Process Theories in Social Psychology*, Guilford, New York.

Cheng, P.W., and Holyoak, K.J., 1985. Pragmatic reasoning schemas. *Cognitive Psychology*, **17**, 391-416.

Cosmides, L., 1989. The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition* **31**, 187-276.

Cummins, D.D., 1996b. Evidence of deontic reasoning in 3- and 4-year-olds. Memory and *Cognition* **24**, 823-829. Italic

Cummins, D.D., Lubart, T., Alksnis, O., and Rist, R., 1991. Conditional reasoning and causation. *Memory and Cognition* **19**, 274-282.

Evans, JStBT, and Over, D., 1996. Rationality and Reasoning. Psychol. Press, Hove, UK.

Evans, JStBT., 2003 In Two Minds: dual process accounts of reasoning. *Trends in Cognitive Science* **7**, 454-9.

Evans, J.St.B.T., Barston, J., and Pollard, P., 1983. On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition* **11**, 295-306.

Fischoff, B., 1982. "Debiasing." In Kahneman, D., Slovic, P. and Tversky, A. (Eds.), *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, New York, pp. 422-444.

Frederick, S., 2005. Cognitive reflection and decision-making. *Journal of Economic Perspectives* **Vol. 19** (4), 25–42.

Gigerenzer, G., Todd, P. M., and the ABC Research Group, 1999. *Simple heuristics that make us smart*, Oxford University Press, New York.

Guthrie, C., Rachlinski , J. and Wistrich, A., 2001. *Inside the Judicial Mind*. Cornell Law Review, **Vol. 86 (4)**.

Hammond, K., 1996. *Human Judgment and Social Policy*, OUP, New York.

Kahneman, D., 2002. 'Maps of bounded rationality: A perspective on intuitive judgment and choice' (Nobel Prize Lecture), accessible at http://www.nobel.se/economics/laureates/2002/kahnemann-lecture.pdf.

Kahneman, D. and Frederick, S., 2005. A Model of Heuristic Judgment. In Holyaoak, K., Morrison, R. (Eds.), *The Cambridge Handbook of Thinking and Reasoning,* CUP, Cambridge, UK, pp. 267-294.

Kahneman, D., Slovic, P., and Tversky, A., 1982. *Judgment under uncertainty: Heuristics and biases,* Cambridge University Press, New York.

Leron, U. and Hazzan, O., 2006. The Rationality Debate: Application of Cognitive Psychology to Mathematics Education. *Educational Studies in Mathematics* **62(2)**, 105-126

Luria, A., 1966. *Higher Cortical Functions in Mind*, Basic Books, New York.

Osman, M., 2004. An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review* **11(6)**, 988-1010.

Pashler, H., 1998. *The psychology of attention*, MIT. Press, Cambridge MA.

Sloman, 1996. The empirical case for the two systems of reasoning. *Psychological Bulletin* **119**, 3-22.

Sloman, 2002. Two Systems of Reasoning, In Gilovich, T., Griffin, D., Kahneman, D. (Eds.), *Heuristics and Biases: the Psychology of Intuitive Judgment,* CUP, Cambridge, UK, pp. 379-98.

Stanovitch, K. and West R., 1998. Cognitive ability and variation in selection task performance. *Thinking and Reasoning* **4**, 193-230.

Tversky, A and Kahneman, D., 1973. Availability: a heuristic for judging frequency and probability. *Cognitive Psychology* **(5)**, 207-37.